TELKOMNIKA, Vol.11, No.1, January 2013, pp. 40~47
ISSN: 2302-4046

Total Variation based Multivariate Shearlet Shrinkage for Image Reconstruction

Chengzhi Deng*, Saifeng Hu, Wei Tian, Min Hu, Yan Li, Shengqian Wang
Nanchang Institute of Technology
No.298 Tianxiang Road, Nanchang City, P.R.China, 330099
*corresponding author, e-mail: dengchengzhi@126.com

Abstract

Shearlet as a new multidirectional and multiscale transform is optimally efficient in representing images containing edges. In this paper, a total variation based multivariate shearlet adaptive shrinkage is proposed for discontinuity-preserving image denoising. The multivariate adaptive threshold is employed to reduce the noise. Projected total variation diffusion is used to suppress the pseudo-Gibbs and shearlet-like artifacts. Numerical experiments from piecewise-smooth to textured images demonstrate that the proposed method can effectively suppress noise and nonsmooth artifacts caused by shearlet transform. Furthermore, it outperforms several existing techniques in terms of structural similarity (SSIM) index, peak signal-to-noise ratio (PSNR) and visual quality.

Keywords*: Shearlet, denoising, total variation, multivariate shrinkage, structural similarity*

1. Introduction

Images are often corrupted by noise during acquisition and transmission, which will lead to significant degradation of image quality for human interpretation and post-processing tasks. The main goal of image denoising is to reduce the noise, while preserving the image features. Partial differential equations and computational harmonic analysis are two widely used classes of methods to achieve this goal.

Wavelet transform as one of the computational harmonic analysis methods has been successfully used in image denoising field. Since Donoho and his coworker pioneered a wavelet denoising scheme by using thresholding [1], there have been hundreds of papers presented to apply or modify the original algorithm. However, when wavelet is used to image denoising, it will lead to oscillatory artifacts along the edges. That is why wavelet fails to capture the geometric regularity along the singularities of surfaces.

In order to overcome this limitation of traditional wavelet, several image representations have been proposed to capture the geometric regularity of a given image. They include curvelet [2], contourlet [3] and bandelet [4]. Recently, Wang and Labate developed a new geometric multiscale transform, named shearlet [5,6] transform which yields nearly optimal approximation properties. Shearlet transform is based on a simple and rigorous mathematical framework which not only provides a more flexible theoretical tool for the geometric representation of multidimensional data, but is also more natural for implementation. In addition, the shearlet can be associated to a multiresolution analysis. Those features ensure that is superior to other transforms in the field of image processing.

Computational harmonic analysis based image denoising always suffer from some oscillations near discontinuities, like nonsmooth edges or pseudo-Gibbs phenomena. Partial differential equations such as total variation (TV) and diffusion-based method are other powerful tools for denoising and can greatly reduce these ringing effects. But cost-heavy computational burden of these methods are not suitable for time-critical application. The TV-synthesized computational harmonic analysis can effectively overcome these problems. Recently, TV minimization combined wavelet [7], complex wavelet [8], ridgelet [9], wave atoms [10], curvelet [11], and shearlet [12] shrinkage have been considered. The hybrid methods show good performance for real application denoising of engineering surfaces and image with complex textures and geometries. However, in those methods the shrinkage thresholds are fixed at first.

Received September 27, 2012; Revised November 20, 2012; Accepted November 27, 2012

The fixed thresholds only consider the individual coefficient magnitudes, but do not take account of the influence of the coefficient distributions.

In this paper, we propose a TV-synthesized multivariate shearlet adaptive shrinkage technique in order to shrink image noise and, at the same time, suppress unwanted nonsmooth artifacts as well. The large shearlet coefficients are adaptively thresholded by the proposed threshold. And the small ones are modified by TV minimization. Section 2 presents the proposed method. Then, in the next section some experimental results are given, which are compared with results of some existing techniques in both SSIM values and visual qualities. Finally, conclusions are drawn .

2. Proposed Method

Supposing we observe noisy image $g = f + n$ where n is independent, zero-mean Gaussian noise. The aim of denoising is to estimate the noise-free coefficient f as accurately as possible according to some criteria. For threshold $\tau \in \Re^+$, define the shrinkage function $S_\tau(x)$ to be $sign(x)(|x| - \tau)$ if $|x| \geq \tau$ and zero otherwise. A denoised estimate $\hat{f}$ from the shearlet transform can be expressed as

$$\hat{f} = \sum_{M_1} \langle g, y_{j,k,l} \rangle y_{j,k,l} + \sum_{M_2} S_t \left(\langle g, y_{j,k,l} \rangle \right) y_{j,k,l} \quad (1)$$

Where M_1 and M_2 represent the indices of the low-frequency coefficients and shearlet coefficients, respectively. Traditional shearlet hard thresholding, which simply sets the small coefficients as zeros, usually bring nonsmooth shearlet-like artifacts. To overcome this drawback, we combined multivariate shearlet adaptive shrinkage and total variation method to improve the performance for image denoising.

2.1. Multivariate adaptive shrinkage

Traditionally, the threshold function is defined as $\tau = c\sigma_n$, where σ_n denotes the standard deviation of white noise, c is a constant and often $c = 3$ or $c = 4$ for different scales. The traditional threshold function only considers the individual shearlet coefficient magnitudes, but dose not take account of the influence of the coefficient models, which is not optimal. In this paper, the models of shearlet coefficient are considered, and the threshold function is adaptively determined by the maximum a posteriori (MAP) estimator.

After shearlet transform, the observed noisy image can be formulated as $\mathbf{y} = \mathbf{x} + \mathbf{n}$. After shearlet transform, the problem can be formulated as $\mathbf{y} = \mathbf{x} + \mathbf{n}$ where $\mathbf{y}$ is the noisy shearlet coefficients, $\mathbf{x}$ is the noise-free shearlet coefficient and $\mathbf{n}$ is noise, which is yet independent, white, zero-mean Gaussian. Using Bayes rule, the maximum a posterior estimator for $\mathbf{x}$ given the noisy observation $\mathbf{y}$ can be easily derived as being

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg\max_{\mathbf{x}} \left\{ p_{\mathbf{x}|\mathbf{y}}(\mathbf{x} \mid \mathbf{y}) \right\} = \arg\max_{\mathbf{x}} \left\{ p_{\mathbf{n}}(\mathbf{y} - \mathbf{x})\, p_{\mathbf{x}}(\mathbf{x}) \right\} \quad (2)$$

where $p_{\mathbf{n}}(\mathbf{y} - \mathbf{x})$ and $p_{\mathbf{x}}(\mathbf{x})$ are the probability density functions of the noise coefficient $\mathbf{n}$ and the noise-free shearlet coefficient $\mathbf{x}$. The Eq.(2) is also equivalent to

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg\max_{\mathbf{x}} \left\{ \log\left[p_{\mathbf{n}}(\mathbf{y} - \mathbf{x}) \right] + \log\left[p_{\mathbf{x}}(\mathbf{x}) \right] \right\} \quad (3)$$

From the equation (3), in order to estimate $\mathbf{x}$, the probability distribution of noise coefficient $\mathbf{n}$ and the prior distribution of noise-free shearlet coefficient $\mathbf{x}$ must be supposed. In this paper, the noise is assumed as additive white Gaussian noise with zero-mean and variance s_n^2. The probability distribution of noise coefficient can be expressed by

$$p_{\mathbf{n}}(\mathbf{n}) = \frac{1}{\left(2p s_n^2\right)^{\frac{d}{2}}} \times \exp\left(- \|\mathbf{n}\|^2 / \left(2 s_n^2\right)\right) \tag{4}$$

where d is the dimension of the noise coefficients vector.

Substituting Eq. (4) into Eq. (3) yields

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg\max_{\mathbf{x}} \left\{ -\frac{d}{2}\log\left(2\pi\sigma_n^2\right) - \frac{\|\mathbf{y}-\mathbf{x}\|^2}{2\pi\sigma_n^2} + \log\left[p_{\mathbf{x}}(\mathbf{x})\right] \right\} \tag{5}$$

Assuming $z(\mathbf{x}) = -\log\left[p_{\mathbf{x}}(\mathbf{x})\right]$ to be strictly convex and differentiable, the following first-order approximation of MAP estimator is always possible

$$\hat{\mathbf{x}} = \mathbf{y} - p s_n^2 z'(\mathbf{y}) \tag{6}$$

where the problem with the estimator in (6) is that the sign of $\hat{\mathbf{x}}$ is often different from the sign of $\mathbf{y}$ even for symmetrical zero-mean densities. Such counterintuitive estimates are possible because ζ' is often discontinuous or even singular at 0, which implies that the first-order approximation is quite inaccurate near 0. To alleviate this problem of "overshrinkage", we use the following modification as the MAP estimator (2) of a nongaussian random variable corrupted by Gaussian noise:

$$\hat{\mathbf{x}} = \begin{cases} sign(\mathbf{y})(|\mathbf{y}| - \boldsymbol{\tau}) & |\mathbf{y}| > \boldsymbol{\tau} \\ 0 & |\mathbf{y}| \le \boldsymbol{\tau} \end{cases} \tag{7}$$

where shrinkage threshold $\boldsymbol{\tau} = p s_n^2 \left| z'(\mathbf{y}) \right|$.

From the equation (7), in order to estimate $\hat{\mathbf{x}}$, the probability distribution of the prior distribution of shearlet coefficient $\mathbf{x}$ must be supposed. It is well known that the statistical models of shearlet coefficients exhibit a sharp peak at zero and heavy tails to both sides of the peak. And due to the multiresolution and multidirection features of shearlet transform, the distribution of its coefficients is a wide range of processes, from heavy-tailed to less heavy-tailed processes. The normal inverse Gaussian (NIG) model is a flexible, four parameter distribution that can describe a wide range shapes. In this paper, we use the following multivariate normal inverse gaussian probability density function to model the noise-free shearlet coefficients. The multivariate normal inverse gaussian distribution of $\mathbf{x}$ is given by

$$p_{\mathbf{x}}(\mathbf{x}) = \frac{d}{2^{(d-1)/2}} \left(\frac{a}{p\, q(\mathbf{x})} \right)^{(d+1)/2} \times \exp\left[p(\mathbf{x})\right] K_{(d+1)/2}\left[a\, q(\mathbf{x})\right] \tag{8}$$

with $p(\mathbf{x}) = d\sqrt{a^2 - b^T \Gamma b} + b^T(\mathbf{x} - m)$ and $q(\mathbf{x}) = \sqrt{d^2 + (\mathbf{x} - m)^T \Gamma^{-1} (\mathbf{x} - m)}$. Where $K_{(d+1)/2}$ denotes the modified Bessel function of second kind with order $(d+1)/2$, $\Gamma \in \Re^{d \times d}$ denotes the correlation matrix. In this letter, we assume the distribution of noise-free shearlet coefficient $\mathbf{x}$ is symmetrical with zero mean this means that $m = 0$ and $b = 0$. Then, from Eq. (8), we have

$$z'(\mathbf{x}) = \frac{\partial \log\left[p_{\mathbf{x}}(\hat{\mathbf{x}})\right]}{\partial \hat{\mathbf{x}}} = \left(\frac{a^2}{d^2 + \mathbf{x}^T \Gamma^{-1} \mathbf{x}} \right)^{1/2} \frac{K_{(3+d)/2}\left[a\sqrt{d^2 + \mathbf{x}^T \Gamma^{-1} \mathbf{x}}\right]}{K_{(1+d)/2}\left[a\sqrt{d^2 + \mathbf{x}^T \Gamma^{-1} \mathbf{x}}\right]} \Gamma^{-1} \mathbf{x} \tag{9}$$

Using [9], we get the adaptive shrinkage threshold $\boldsymbol{\tau}$.

$$\boldsymbol{\tau} = ps_n^2 \left| z\left(\mathbf{y}\right) \right| = \left(\frac{a^2}{d^2 + \mathbf{y}^T \mathbf{G}^{-1} \mathbf{y}} \right)^{1/2} \frac{ps_n^2 K_{(3+d)/2}\left(a\sqrt{d^2 + \mathbf{y}^T \mathbf{G}^{-1} \mathbf{y}} \right)}{K_{(1+d)/2}\left(a\sqrt{d^2 + \mathbf{y}^T \mathbf{G}^{-1} \mathbf{y}} \right)} \mathbf{G}^{-1} \mathbf{y} \tag{10}$$

where $K_d(\cdot)$ is the modified Bessel function of the second kind with index d.

2.2. TV combined multivariate adaptive shrinkage

TV minimization is adopted to shrink the shearlet coefficients which have been set to zero by the multivariate adaptive shrinkage procedure (7). For a function $f(x): \Re^2 \supseteq \Omega \to \Re$ with $|\nabla f(x)| \in L^1(\Re^2)$, the TV function is defined as

$$\mathrm{TV}(f) = \int_\Omega |\nabla f(x)| dx \tag{11}$$

To circumvent computational difficulties arising from the non-differentiability of the modulus at zero, the TV function is often replaced by

$$\mathrm{TV}(f) = \int_\Omega \sqrt{|\nabla f(x)|^2 + \beta^2}\, dx \tag{12}$$

where β is a small positive parameter. In this paper, the β is set to be 10^{-4}. The discrete version of the TV function for $f = (f_{i,j})_{(i,j)\in I_n^2}$ is given by

$$\mathrm{TV}(f) = \sum_{i,j} \sqrt{\left|(\delta_1 f)_{i,j}\right|^2 + \left|(\delta_2 f)_{i,j}\right|^2 + \beta^2} \tag{13}$$

where $(\delta_1 f)_{i,j} = f_{i+1,j} - f_{i,j}$, $(\delta_2 f)_{i,j} = f_{i,j+1} - f_{i,j}$. Then our proposed model can be formulated as follows.

$$\begin{cases} \arg\min_{u \in U} \{\mathrm{TV}(u)\} \\ U = \left\{ u : u \in V_B(\Omega), u = \sum_{M_1} \langle f, y_{j,k,l} \rangle y_{j,k,l} + \sum_{M_2^c} \langle f, y_{j,k,l} \rangle y_{j,k,l} \right\} \end{cases} \tag{14}$$

where V_B denotes bounded variation. M_1 denotes the indices of the low-frequency shearlet coefficients. M_2^c denotes the indices of shearlet coefficients which are greater than the proposed shrinkage threshold τ. Since the function (13) is convex, problem (14) has a solution. In this paper, a projected gradient descent scheme is adopted to diminish the TV norm. Define $g_{\mathrm{TV}}(f^\ell)$ to be the subgradient of $\mathrm{TV}(f^\ell)$, and then, a solution of (14) can be computed by following.

$$f^{\ell+1} = f^\ell - t_\ell \mathrm{P_V}\left(g_{\mathrm{TV}}\left(f^\ell\right)\right) \tag{15}$$

where $\mathrm{P_V}$ denotes the orthogonal projection onto the subspace where shearlet coefficients are smaller than the proposed shrinkage threshold $\boldsymbol{\tau}$.

2.3. Implementation of proposed method

Inputting noisy image f, $\beta = 10^{-4}$, and time step size t_ℓ. The detailed procedure to perform the total variation based shearlet adaptive shrinkage for image denoising is described as follows.

1. Perform discrete shearlet transform to obtain low-frequency coefficients $C_{j_0,k,\ell}(f)$ and shearlet coefficients $d_{j,k,\ell}(f)$
2. Apply the proposed multivariate adaptive threshold (10) and the threshold function (7) to shrink the shearlet coefficients.
3. Compute initial reconstruction f^0 by inverse shearlet transform of the shrinkage shearlet coefficients.
4. Minimize the TV norm of f^ℓ by the following steps for $\ell = 1,2,...$:

   4.1. Compute the subgradient $g_{\mathrm{TV}}(f^\ell)$ of f^ℓ.

   4.2. Compute the discrete shearlet coefficients $C_{j_0,k,\ell}(g_{\mathrm{TV}}(f^\ell))$, $d_{j,k,\ell}(g_{\mathrm{TV}}(f^\ell))$.

   4.3 Compute the orthogonal projection $\mathrm{P_V}(g_{\mathrm{TV}}(f^\ell))$ and set $f^{\ell+1} = f^\ell - t_\ell \mathrm{P_V}(g_{\mathrm{TV}}(f^\ell))$.

3. Results and Analysis

In this section, we present results of our proposed algorithm and compare them with some of the recent existing techniques, namely BLS-GSM [13] with non-decimated wavelet transform, Shearlet hard shrinkage (SH-H), and total variation based shearlet hard shrinkage (TV-SH-H) [12]. Extensive experiments are conducted on three standard grayscale test images with distinctly different features, corrupted by simulated additive Gaussian white noise at different power levels. And we use 20 iterations and 0.2 step size for the computation of TV minimization in all experiments.

In this paper, we chose SSIM and PSNR as the objective evaluation criterion. Although PSNR can measure the difference between two images, it is well-known that it may fail to describe the visual perception quality of an image. The PSNR is not the best choice for image quality assessment (IQA). Among the recently proposed IQA approaches, the structural similarity (SSIM) index [14] has emerged as a promising measure for image visual quality assessment. Compared with PSNR, SSIM can better reflect the structure similarity between target image and the reference image. The SSIM and PSNR values for implementations using different images and different noise levels are list in Table 1. From Table 1, it is seen that the proposed method consistently gives a larger value of SSIM and PSNR compared to the other methods, which indicating a better preservation of structure in the denoised images,especially to the textured images (such as Barbara).

Table 1. SSIM and PSNR comparison of some grayscale image denoising results

Image	Lena (512×512)			Barbara (512×512)			Boat (512×512)		
Noise std	10	20	30	10	20	30	10	20	30
				SSIM comparison					
BLS-GSM	0.896	0.836	0.790	0.916	0.834	0.759	0.867	0.781	0.716
SH-H	0.906	0.865	0.826	0.929	0.869	0.808	0.871	0.803	0.746
TV-SH-H	0.906	0.867	0.833	0.929	0.870	0.810	0.871	0.804	0.749
Proposed	**0.907**	**0.871**	**0.839**	**0.933**	**0.878**	**0.821**	**0.879**	**0.812**	**0.753**
				PSNR comparison (in dB)					
BLS-GSM	34.64	31.42	29.52	33.13	29.03	26.89	32.84	29.42	27.52
SH-H	35.09	32.36	30.63	33.75	30.09	27.87	32.76	30.01	28.26
TV-SH-H	35.18	32.42	30.69	34.06	30.23	27.95	32.99	30.07	28.31
Proposed	**35.27**	**32.47**	**30.71**	**34.47**	**30.64**	**28.35**	**33.28**	**30.23**	**28.36**

For visual evaluation, one example using the standard “Lena” and “Barbara” image with noise level 40 is given. Figure 1 and figure 2 show denoising results on a cropped subregion of Barbara and Lena. Form the figures we can seen that nonsmoothness oscillations along the

edges can be seen obviously in wavelet domain (BLS-GSM) due to the poor ability of wavelet at presenting curve singularities. Due to the sparse representation of shearlet transform for curve singularities, the denoising methods in shearlet domain show good performance for the edge preserving denoising. From Figure 1(d) and 2(d), we can see that noise is removed more effectively than methods in wavelet domain. However, it suffers from the shearlet-like artifacts. That is why SH-H method simply sets the small shearlet coefficients to be zero. TV minimization is also a noise shrinkage way. It can greatly reduce ringing effects. But too many details will be smoothed out. TV-SH-H method use TV minimization to adjust the small shearlet coefficients and fix the significant coefficients unchanged. The results are shown in Figure 1(e) and 2(e), where the shearlet-like artifacts are reduced. But there are still some noise left, particularly around discontinuities. In this paper, adaptively thresholding the significant coeffieients and adjusting the small shearlet coefficients by TV minimization are effective in denoising and recovering edges while the shearlet-like artifacts are greatly suppressed. The results are shown in Figure 1(f) and 2(f).

Figure 1. Denoising of corrupted image Barbara ($\sigma_n = 40$). (a) The original image. (b) Noisy image. (c) Output from BLS-GSM. (d) Output from SH-H. (e) Output from TV-SH-H. (f) Output from proposed method

The differences between the noisy and denoised Lena images obtained by the four denoising schemes are shown in Figure 3. Figures 4(a)-(d) present suppressed components by BLS-GSM, SH-H, TV-SH-H and the proposed method. From Figures 4(a), it is clear that BLS-GSM method remove the noise, meanwhile the image features are lost. Compared with wavelet method (BLS-GSM), methods based shearlet (SH-H, TV-SH-H) can reduce noise more efficiently. But there still have image features losted. From Figure 4(d), it is shown that the proposed scheme removes mostly noise with less image features, thus indicating a better suppression of noise and retention of image features in the denoisied image.

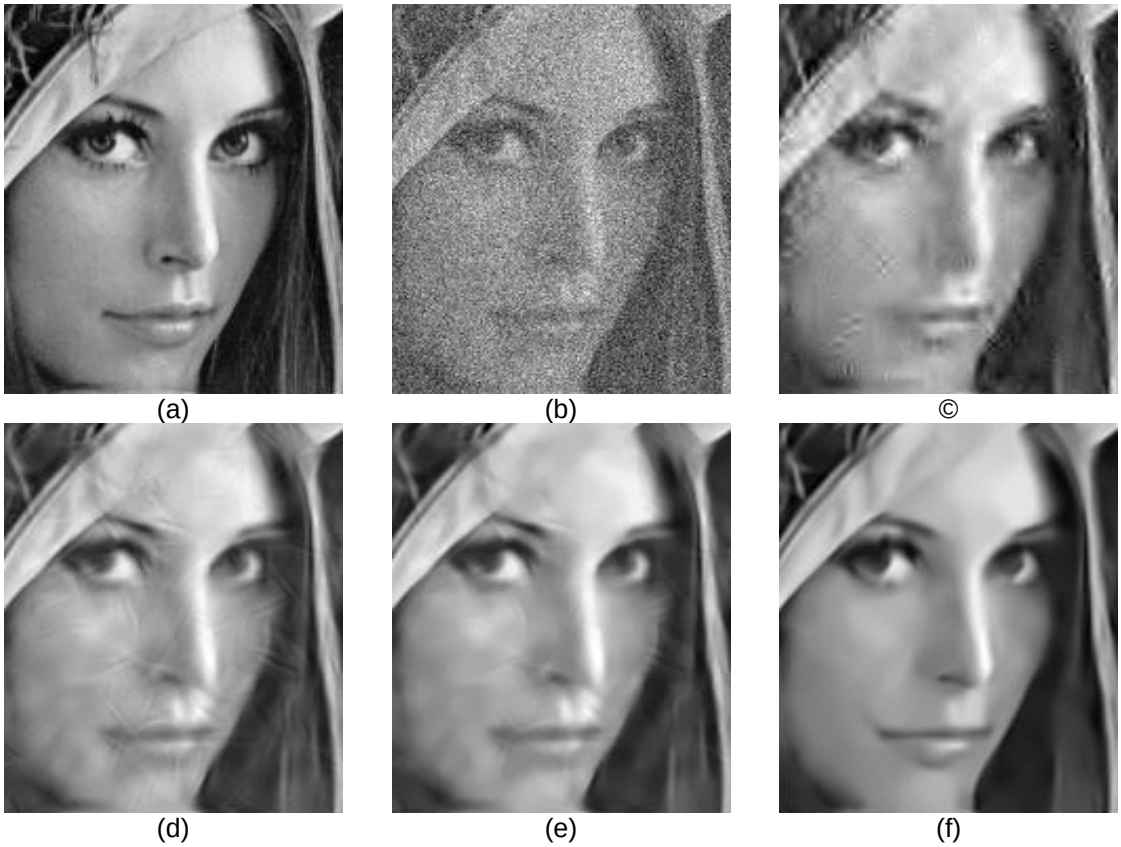

Figure 2. Denoising of corrupted image Lena ($\sigma_n = 40$). (a) The original image. (b) Noisy image. (c) Output from BLS-GSM. (d) Output from SH-H. (e) Output from TV-SH-H. (f) Output from proposed method

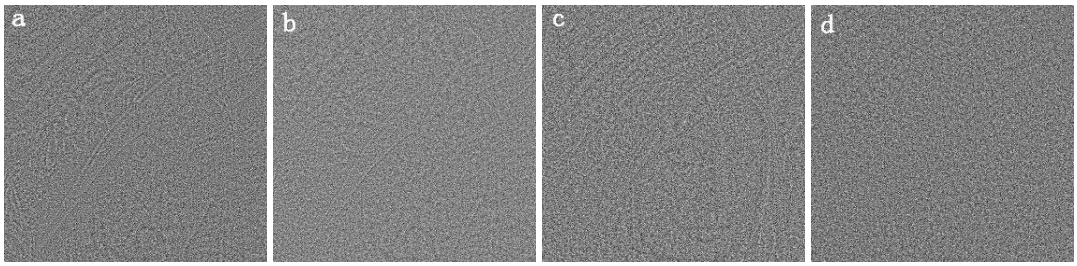

Figure 3. Removed components of Lena ($\sigma_n = 40$) by different schemes. (a) BLS-GSM. (b) SH-H. (c) TV-SH-H. (d) the proposed scheme

4. Conclusion

In this paper, an efficient algorithm is proposed for removing noise from corrupted image by incorporating a total variation based multivariate shearlet adaptive shrinkage. A multivariate adaptive threshold is firstly used to reduce the noise. And then, TV minimization is employed to modify thresholded shearlet coefficients in order to suppress nonsmooth artifacts near the edges and the shearlet-like artifacts. To demonstrate the superior performance of the proposed method, extensive experiments have been conducted on several standard test images. Experimental results show that the proposed method achieves state-of-the-art performance both visually and in terms of SSIM and PSNR.

5. Acknowledgement

The authors would like to thank the anonymous reviewers for their helpful comments. The work was jointly supported by the National Natural Science Funds of PR China under grant 61162022, Natural Science Funds of Jiangxi P.R. China under grant 2009GZW0020 and 2010GZW0049.

References

[1] Donoho DL, Johnstone IM. *Ideal spatial adaptation via wavelet shrinkage*. Biometrika. 1994; 81(2): 425-355.
[2] Candès EJ, Demanet L, Donoho DL, Ying L. *Fast discrete curvelet transforms*. Multiscale Modeling Simulation. 2006; 5(3): 861-899.
[3] Do MN, Vetterli M. The contourlet transform: An efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*. 2005; 14(2): 2091-2106.
[4] Mallat S, Peyre G. *A review of bandlet methods for geometrical image representation*. Numerical Algorithm. 2007; 44(2): 205-234.
[5] Easley GR, Labate D., Lim W. *Sparse directional image representations using the discrete shearlet transform*. Applied and Computational Harmonic Analysis. 2008; 25(1): 25-46.
[6] Lim W. The Discrete Shearlet Transform: A New Directional Transform and Compactly Supported Shearlet Frames. *IEEE Transactions on Image Processing*. 2010; 19(5): 1166-1180.
[7] Durand S, Froment J. Reconstruction of wavelet coefficients using total variation minimization. *SIAM Journal on Scientific Computing*. 2003; 24(10): 1754-1767.
[8] Ma J. Towards artifact-free characterization of surface topography using complex wavelets and total variation minimization. *Applied Mathematics and Computation*. 2005; 170(2): 1014-1030.
[9] Ma J, Fenn M. Combined complex ridgelet shrinkage and total variation minimization. *SIAM Journal on Scientific Computing*. 2006; 28(3): 984-1000.
[10] Plonka G, Ma J. Nonlinear regularized reaction-diffusion filters for denoising of images with textures. *IEEE Transactions on Image processing*.2008; 17(8): 1283-1294.
[11] Ma J, Plonka G. Combined curvelet shrinkage and nonlinear anisotropic diffusion. *IEEE Transactions on Image processing*. 2007; 16(9): 2198-2206.
[12] Easley GR, Labate D, Colonna F. Shearlet-based total variation diffusion for denoising. *IEEE Transactions on Image processing*. 2009; 18(2): 260-268.
[13] Portilla J, Strela V, Wainwright M, Simoncelli P. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Transactions on Image Processing*. 2003; 12(11): 1338-1351.
[14] Wang Z, Bovik AC. Mean squared error: love it or leave it? - A new look at signal fidelity measures. *IEEE Signal Processing Magazine*. 2009; 26(1): 98-117.